CEAM: The Effectiveness of Cyclic and Ephemeral Attention Models of User Behavior on Social Platforms

Farhan Asif Chowdhury[1], Yozen Liu[2], Koustuv Saha[3], Nicholas Vincent[4], Leonardo Neves[2], Neil Shah[2], Maarten W. Bos[2]

[1]University of New Mexico, [2]Snap Inc., [3]Georgia Tech, [4] Northwestern University
[1]fasifchowdhury@unm.edu, [2]{yliu2, lneves, nshah, mbos}@snap.com,
[3]koustuv.saha@gatech.edu, [4]nickvincent@u.northwestern.edu

Abstract

To improve the user experience as well as business outcomes, social platforms aim to predict user behavior. To this end, recurrent models are often used to predict a user's next behavior based on their most recent behavior. However, people have habits and routines, making it plausible to predict their behavior from more than just their most recent activity. Our work focuses on the interplay between ephemeral and cyclical components of user behaviors. By utilizing user activity data from social platform Snapchat, we uncover cyclic and ephemeral usage patterns on a per user-level. Based on our findings, we imbued recurrent models with awareness: we augment an RNN with a cyclic module to complement traditional RNNs that model ephemeral behaviors and allow a flexible weighting of the two for the prediction task. We conducted extensive experiments to evaluate our model's performance on four user behavior prediction tasks on the Snapchat platform. We achieve improved results on each task compared against existing methods, using this simple, but important insight in user behavior: both cyclical and ephemeral components matter. We show that in some situations and for some people, ephemeral components may be more helpful for predicting behavior, while for others and in other situations, cyclical components may carry more weight.

1 Introduction

Online social platforms such as Facebook, Twitter, and Snapchat have become a ubiquitous and integral part of our daily lives to communicate and interact with others, find entertainment, and learn about current events. Content creators, providers, and consumers benefit when the right content is served to the right person at the right time. Additionally, better content recommendation enhances user engagement and interest on the platforms. Therefore, these social platforms proactively customize and personalize content, interactive features, and ads based on a user's behavior and preferences, to improve user satisfaction and engagement (Berkovsky, Kuflik, and Ricci 2008). A prerequisite for successful personalization is predicting user behavior ahead of time. However, human behavior is complex, and is confounded by several on-platform and off-platform factors (Banovic et al. 2016). Those factors include time of the day (e.g. shorter chat sessions during lunch time vs. longer chat sessions in the evening) (Golder, Wilkinson, and Huberman 2007), day of the week (binge content-watching on weekends), momentary emotional state (exploring contents when bored) (Kapoor et al. 2015a), social presence (sharing photos when meeting with friends), and other factors (Habib, Shah, and Vaish 2019; Koren 2009; Cheng et al. 2017).

Recent works on modeling user behavior have utilized user activities in the recent past to make a prediction for an upcoming session (Hidasi et al. 2015; Jing and Smola 2017; Beutel et al. 2018; Zhu et al. 2017; Saha et al. 2021b). Hidasi et al. (Hidasi et al. 2015) utilized Recurrent Neural Networks (RNNs) to model the dynamics on sequential behaviors. Jing and Smola incorporate timing of the sessions and time-interval between consecutive sessions in a recurrent model. A few recent works have shown better performance by using Long Short-Term Memory (LSTM) networks with time-gates (Zhu et al. 2017), and exploiting contextual information, such as location and device (Beutel et al. 2018; Tang et al. 2020). These methods exploit the recency or ephemerality in user behavior to dynamically model the temporal variation in online behavior.

So behaviors do vary based on time, and predicting behavior from recent behaviors has proven valuable. There is, however, another insight we can incorporate: as humans, we imbibe into habits of what we do, how we do, and when we do; arguably these habits make humans predictable. Habits and routines are — among other drivers — driven by seasons, circadian rhythms (Murnane et al. 2015), and school, work, and workout schedules (Kooti et al. 2017). Given that our offline and online behaviors are often intertwined, our online activities exhibit daily, weekly, and monthly cycles (Saha et al. 2021a). Previous works have demonstrated such regularities on multiple platforms (Grinberg et al. 2013; Zhang et al. 2013). Golder, Wilkinson, and Huberman (Golder, Wilkinson, and Huberman 2007) demonstrated consistent weekly and seasonal patterns of social interaction among college students on Facebook. Grinberg et al. (Grinberg et al. 2013) identified daily and weekly patterns of food consumption and nightlife activity using Foursquare check-in data. A few methods have utilized cyclicity by incorporating session-timing as contextual information to capture the temporal dependency of user behavior (Beutel et al. 2018; Jing and Smola 2017). However, using time as a context with a fixed effect for all users limits the predictive models from learning user-specific temporal patterns. While this is probably less an issue for circadian rhythms and seasonality effects, it may leave user-specific variance unexplained.

In this work, we argue that existing methods cannot fully exploit the cyclicity in user behavior, and through effective utilization, better predictive performance can be achieved. Moreover, although cyclicity is well-understood in global engagement and platform metrics (Zhang et al. 2013; Golder and Macy 2011), it is less understood on a per user-level. Therefore, to address these two limitations, we guide our work based on the following research questions:

- **RQ1**: Does individual user behavior on social platforms exhibit cyclical properties, and does cyclicity vary across users?
- **RQ2**: To what extent can user behavior prediction on a social platform be improved by exploiting per user-level cyclicity?

To answer the first research question, we aimed to uncover the temporal dynamics of individual user-level behavior by analyzing user activity from Snapchat, a popular multimedia-driven online social platform. We demonstrated empirically that user behavior is largely driven by regularities (cyclicity) and ephemeral actions. We observed similar temporal variations of user activity level across user cohorts on a daily and weekly basis. While examining regularities at the individual user level, we also noticed varying levels of cyclicity across users. The variations of temporal activity level across user cohorts and the difference of cyclicity across users signify the necessity of modeling cyclical temporal dynamics, at an individual user level.

To answer the second research question, we aimed to model cyclic and ephemeral behaviors jointly, by utilizing a novel end-to-end neural framework which builds on well-validated recurrent networks. We augmented the traditional, ephemeral LSTM module with an additional LSTM head that utilizes historical activity data over a longer period to capture individual user level cyclicity. Specifically, we aggregated user activities in a particular time-frame (i.e., hour of the day) to capture cyclicity in the corresponding prediction time-frame. We further employed an attention mechanism to adaptively fuse information from both heads. We evaluated the performance of our model on two anonymized user activity datasets collected from Snapchat. We defined four behavior prediction tasks that are generalizable to other platforms and compared the performance against existing baseline methods to demonstrate that the simple addition of cyclicity modeling can effectively lead to improved accuracy from baselines on all four prediction tasks.

Through ablation studies, we examined the impact of each module in our model. We also conducted post-hoc sanity analysis on improvements over users with higher cyclicity, attention on cyclic and ephemeral module when various amounts of data are available, and the performance at different times of the day. To summarize, our contributions are:

- We show that user behavior on social platforms is driven by cyclicity and ephemerality, and the patterns and level of cyclicity vary across users.
- We leverage per user-level cyclicity by adding a cyclic-LSTM module along with the existing ephemeral-LSTM architecture to jointly model cyclicity and ephemerality through attention-based adaptive fusion.
- We leverage regularities in user-behavior to achieve personalization without using any personally identifiable information, which is a timely approach considering societal concerns about data breaches.
- We outperform existing methods by on average 7% (up to 10%) *macro f1-score* on four user behavior prediction tasks using two real-world datasets from Snapchat. We demonstrate that our model can effectively model cyclicity.

2 Related Works

User Behavior Modeling. Since the ever-increasing popularity of social platforms, many studies have sought to understand, characterize and model people's usage of these platforms. Such efforts include user engagement prediction, user churn rate prediction, user intention prediction, among others (Tang et al. 2020; Kawale, Pal, and Srivastava 2009; Yang et al. 2018; Liu et al. 2019; Verbeke, Martens, and Baesens 2014; Lo, Frankowski, and Leskovec 2016). However, the majority of these works focus on aggregated user activities in the context of a long period (or window) of time and not on a session-level. Recent research revealed the potential of accounting for person-centered modeling on user behaviors (Das Swain et al. 2019; Saha et al. 2021a). Related to our problem space (Kooti et al. 2017) utilized first-minute user activities in a Facebook session to predict the activity duration for the rest of the session. Another closely related work (Kurashima, Althoff, and Leskovec 2018) modeled the action logging of mobile health app to predict the next action based on the history of actions. This work formulated a probabilistic temporal point process model that considers temporal variation, short-term dependency, and long term periodic effect. However, they model time-varying action propensity on a global level rather than on an individual level.

Regularities in Online Behavior. Several previous works have explored regularities in online and offline human behavior and identified daily, weekly, monthly, and seasonal patterns (Golder and Macy 2011; Culotta 2014; Murnane et al. 2015). Golder and Macy identified diurnal and seasonal patterns of individual mood based on Twitter posts (Golder and Macy 2011). Golder, Wilkinson, and Huberman found consistent weekly and seasonal patterns of social interaction among college students on Facebook (Golder, Wilkinson, and Huberman 2007). Grinberg et al.(Grinberg et al. 2013) show daily and weekly patterns of eating, drinking, shopping, and nightlife in human behavior using Foursquare checkins. Moreover, it has been shown that people tend to reply to emails faster in the mornings and on weekdays (Kooti et al. 2015). Pierson, Althoff, and Leskovec proposed a Cyclic Hidden Markov Model to detect and model cycles in human menstrual cycle symptoms and physical activity tracking data (Pierson, Althoff, and Leskovec 2018). A few other works have studied regularities in the context of repeated actions, for example repeated web search queries (Teevan et al. 2006), web page revisitation patterns (Adar, Teevan, and

Dumais 2008), music listening (Kapoor et al. 2015b), and video binge watching (Trouleau et al. 2016). Recently, Saha et al. causally examined the effectiveness timing ads based on person-centered contextualized modeling of user behavior on online platforms (Saha et al. 2021b). Several recent works have proposed models that exploit the recency or ephemerality of user behavior (Anderson et al. 2014; Benson, Kumar, and Tomkins 2016; Kapoor et al. 2015b). Anderson et al. modeled the dynamics of repeat consumption based on recency. In contrast, user-level cyclicity has not been exploited for user behavior modeling.

Recurrent Models for Behavior Modeling. In the recent past, recurrent models have shown promising results in a multitude of user behavior modeling tasks, mostly in the context of recommendation systems (Hidasi et al. 2015, 2016; Smirnova and Vasile 2017) with applications in next basket (or item) recommendation (Jing and Smola 2017), streaming content recommendation (Beutel et al. 2018), check-in location prediction (Chen et al. 2018) etc. Jing and Smola utilized the session time, time interval between sessions and contextual information as features to improve performance. Zhu et al. introduced time-gate for LSTM to model time intervals between sessions to improve the recommendation performance. Beutel et al. improved the contribution of contextual features by using second order neural network to directly modify the neural network hidden states. Although these recurrent models consider recent temporal dynamics and contextual information, they are limited in their capability to capture long term cyclical effects. Moreover, the majority of these methods learn user embeddings for each user to incorporate user-centric features, which is both static and privacy intrusive. These models are also limited in functionality for new users in cold start situations.

3 Task Description

We consider a general social platform where each user u represents a registered user. Each user can engage with the platform by using several in-platform features, such as chatting with a friend in Snapchat, posting a tweet on Twitter, watching video clips on Facebook, or reacting to photos on Instagram; we call these user activities a. A session s consists of a continuous sequence of activities, and two sessions are separated by more than a specific time interval. Each session is represented by a feature vector f, which contains each activity's aggregated amounts in that session. For example, the number of photos shared and the number of videos watched appear in f.

We formulate a user activity prediction task where, for each user, we have a sequence of previous sessions consisting of session-activity features along with the timing of the session. Let us consider a set of users U, and each user $u \in U$ has a sequence of historical sessions $H_u = H_{u_1}, H_{u_2}, ...H_{u_N}$, where $H_{u_i} = \{(f,t)\}$. f represents the activity feature vector, and t represents the timing of that session. Here, $t < t_I$, t_I = timing of the session to be predicted. Our user behavior prediction task can be formalized as follows:

Problem 1 (User Behavior Prediction) *Given a set of users U and sequences of historical session H; for an upcoming session s at time t of user u, predict the amount m of user activity a.*

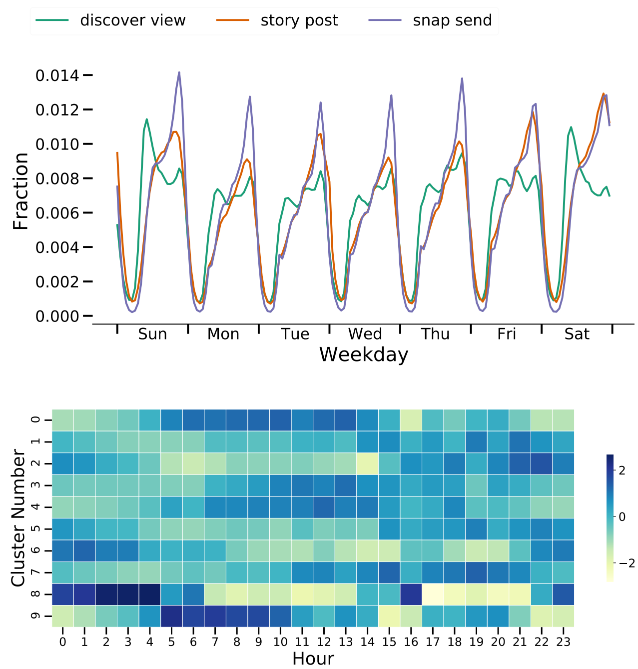


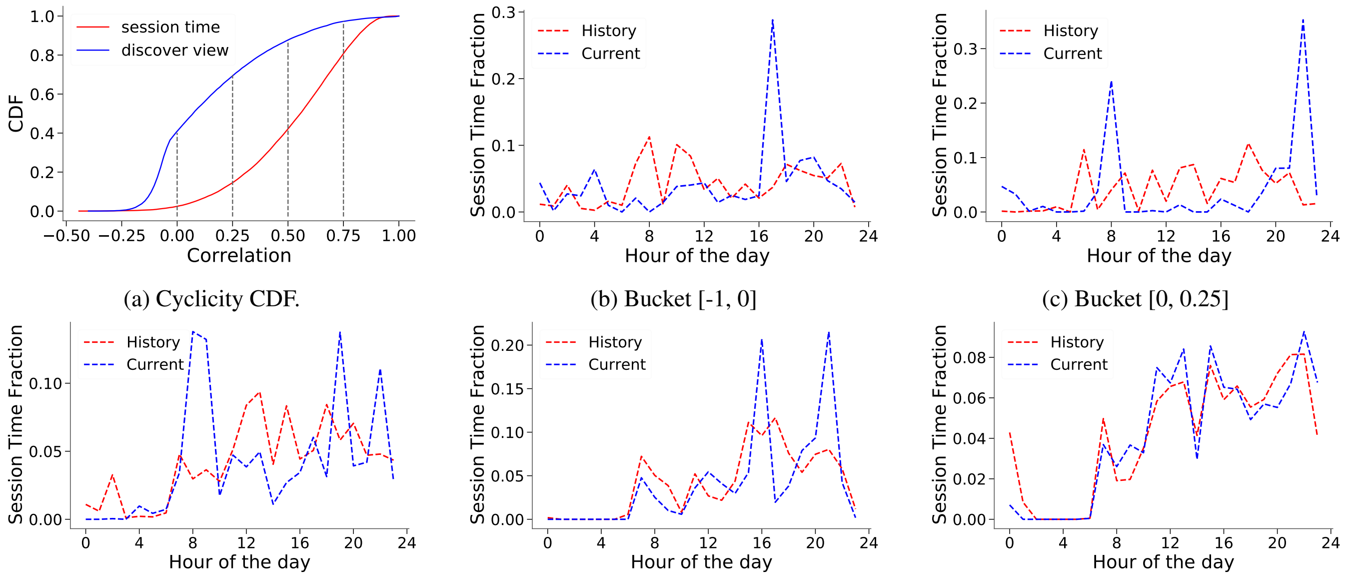
Figure 1: (top) Diurnal cyclicity in global user activity. The y-axis represents fraction of activity that happens in each weekday-hour over a single week, aggregated across all users. (bottom) DISCOVER VIEW activity across user clusters by hour (values are z-transformed), darker shades indicate higher activity amount. Here, users are clustered based on the aggregated value of their Snapchat activities (such as frequency and amount of communication and content consumption).

4 RQ1: Temporal Dynamics and Cyclicity in User Behavior

In this section, we investigate two specific aspects of user behavior (1) routine or *cyclic* behavior, (2) transient or *ephemeral* behavior.

Dataset Description

To examine dynamics and patterns of user behavior in social platforms, we conduct our study on an anonymized user activity dataset from the smartphone app Snapchat, a popular social, multimedia, instant messaging platform used by more than 230M users worldwide (Statista 2020). First, we randomly sample 20K users who were active at least once in each month over the span of seven weeks from January 6, 2020 to February 23, 2020. Subsequently, we collect longitudinal user activity data for these users in the same period. The longitudinal nature of data allows us to study each user's on-platform activity spread across several "sessions" of participation on Snapchat. Our study defines a session to start



(a) Cyclicity CDF. (b) Bucket [-1, 0] (c) Bucket [0, 0.25]

(d) Bucket [0.25, 0.50] (e) Bucket [0.50, 0.75] (f) Bucket[0.75, 1]

Figure 2: (a) Shows the cyclicity CDF; here, most users are cyclic, with more than 80% of users having a cyclicity score of more than 0.25. (b)-(f) show hourly average activity (session duration) in the first three weeks (history) and fourth week (current) for five example users, each selected from one of the correlation value buckets ([-1, 0], [0, 0.25], [0.25, 0.50], [0.50, 0.75], [0.75, 1]), as designated with dashed lines in (a). Here, corresponding with the increasing cyclicity score from (b)-(f), the similarities between historical and current activity increases, too.

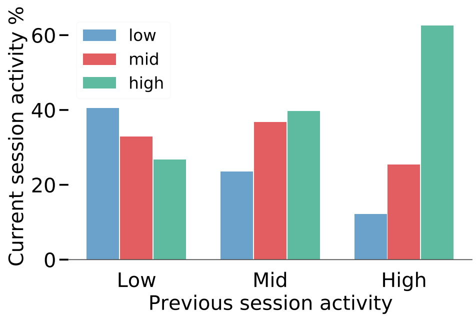
when a user opens the app, and a session to end when they close the app or they leave the app inactive for 15 seconds. We use the first four weeks of data for making empirical observations to motivate our modeling approach. Later (in Section 6), we use these four weeks of data to train our predictive models, and hold-out the subsequent three weeks to evaluate our models.

Cyclicity

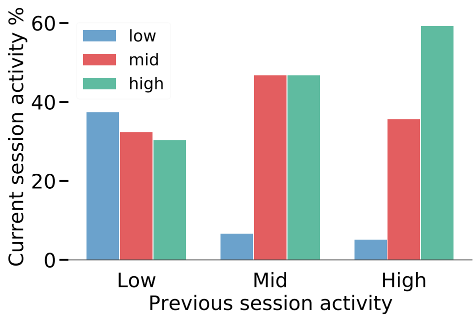
Human behavior is largely affected and controlled by circadian rhythms, sleep habits, work, and leisure schedules (Kooti et al. 2017), which is also reflected in online behaviors (Murnane et al. 2015). Consequently, people's activities on online social platforms often show diurnal and weekly cycles (Golder, Wilkinson, and Huberman 2007). To explore regularities at an aggregated level, in Figure 1 (top), we show the global usage pattern for three major Snapchat in-app activities (DISCOVER VIEW, STORY POST, SNAP SEND) for all users in our dataset. We can observe obvious daily patterns and notable hourly variation in all three activity types. The observed hourly variation in activity level is highly aligned with the human circadian rhythm and work-leisure schedule. Both STORY POST and SNAP SEND activities get higher traction as the day progresses and work-to-leisure transition happens. Additionally, DISCOVER VIEW shows irregularity on weekends with spikes in the mornings that correspond to leisure hours.

Next, we explore temporal patterns at the user cohort level. First, we use k-means clustering ($k = 10$) to cluster users based on the aggregated value of their Snapchat activities (such as frequency and amount of communication and content consumption). We use the Elbow heuristic (Satopaa et al. 2011) to select the optimal cluster numbers. To visually examine the temporal variability of activities across clusters, in Figure 1(bottom), we plot a heatmap of mean aggregated DISCOVER VIEW activity (z normalized) for each cluster over the hours, where the vertical axis represents each cluster of users. Here, we observe that DISCOVER VIEW activity varies strongly both across and within clusters over the hours. We observed similar distinct usage patterns across user cohorts for several other app activities.

Intrigued by the global cyclicity and temporal variations across cohorts, we further investigate cyclicity on a per-user level. To quantify individual user's cyclical usage patterns, we define a metric, *cyclicity*, that captures regularities in user behavior. We operationalize the metric per user *cyclicity* by assigning each user a score, defined as the Pearson correlation of users' average hourly activity level for the first three weeks and their average hourly activity level in the fourth week.In Figure 2a, we plot the cumulative distribution functions (CDFs) of *cyclicity* calculated based on SESSION TIME and DISCOVER VIEW. We notice that user behavior is more cyclic based on SESSION TIME than DISCOVER VIEW, which is intuitive as the former encompasses all forms of in-app activities contrary to one specific use case with the latter. From here on, *cyclicity* refers to *cyclicity* calculated with session-time, unless otherwise specified. Overall, most users display a certain extent of *cyclicity*, where 80% of users have a cyclicity score of more than 0.25. Notably, 25% of users have a cyclicity score of more than 0.75. As motivat-



(a) SESSION TIME



(b) DISCOVER VIEW

Figure 3: We show the correlation of user's activity level between consecutive sessions. We categorize each session into three categories based on the level of activity. We show the distribution of current session's activity level in relation to the previous session's activity level for two cases; (a) SESSION TIME, (b) DISCOVER VIEW. In both cases, consecutive sessions are more likely to belong to the same category.

ing examples of varying levels of cyclicity across users, in Figure 2b-2f, we show hourly average SESSION TIME in the first three weeks (history) and fourth week (current) for five randomly sampled users from each bucket.

Ephemerality

Several previous works have demonstrated the continuity of human action and intent in online behavior after a short interval (Zhang et al. 2013; Teevan et al. 2007). For example, Zhang et al. (Zhang et al. 2013) showed that a session's length can be an indicator of successive session length on the popular music streaming platform Spotify. To infer this causal relation, they correlated the length of two successive sessions. Similarly, to explore ephemerality on social platforms, we correlate consecutive session length in Snapchat. First, we divide all sessions into three categories (low, mid, and high) based on session duration. We select the session duration thresholds to roughly have a similar number of sessions in each category. Afterward, we calculate the distribution of each category in relation to the previous session's category. In Figure 3a, in the percentage distribution plot, we notice that for all three categories, their likelihood is highest when the previous session is similar.

To further concretize the notion of ephemerality, we perform similar exploration on another core Snapchat activity, DISCOVER VIEW. As before, we quantify the DISCOVER VIEW activity sessions into three classes with equal class proportions. In Figure 3b, we show the percentage distribution of categorized DISCOVER VIEW depending on the previous session's category. Again, we observe that the successive session's DISCOVER VIEW activity level is similarly correlated as it is for session length. Although these simplified explorations do not consider the more complex dynamics of ephemerality, e.g., the interplay of multiple consecutive sessions or the effect of time-interval, as showcased in prior related studies (Zhang et al. 2013), they can well justify the presence of recency or ephemerality on a social platform in an interpretable way.

Summary findings. (1) User behavior in Snapchat shows strong daily patterns and continuity of activities in consecutive sessions. (2) We observe distinct temporal patterns across user cohorts. (3) We notice that users are in general cyclic; however, the cyclicity level varies across users. These findings motivate us towards our ensuing analyses where we target to better predict user behavior in social platforms by leveraging per user-level cyclicity.

5 RQ2: Jointly Leveraging Cyclicity & Ephemerality

In this section, we formulate a user behavior prediction framework on social platforms to answer the second research question. To improve user behavior prediction by leveraging per user-level cyclicity, we seek to model user behavior in a session as a function of the user's recent activity and user's historical activity around a particular time period (i.e., hour of the day, day of the week). In a modular approach, we model the user's short-term ephemeral behavior and long-term cyclic behavior separately by two independent modules, termed as *ephemeral module* and *cyclic module* respectively, and fuse both for final prediction in the end. Moreover, traditional user behavior models incorporate personalization by utilizing user-typographic and demographic information, which can be privacy-intrusive, biased, and exclusionary. By exploiting the regularities in individual user behavior, we achieve personalization in a privacy-preserving fashion, as it only requires individual user's cyclic activity history to learn user-specific temporal preferences. Therefore, our method does not require any user-centric information, identifier, or demographic data, which is a key strength of our method, to which we return in the Discussion.

Ephemeral Module. Prior works on modeling sequential user behavior data have shown the superiority of the Recurrent Neural Networks (RNNs) over latent variable models (i.e., hidden Markov model) in capturing the short-term temporal dynamics of user behavior (Hidasi et al. 2015; Jing and Smola 2017). Inspired by these successful use cases, we utilize an RNN to model the ephemerality in user behavior. In particular, we use the Long Short-Term Memory (LSTM) network, an improved variant of traditional RNNs that addresses the vanishing gradient problem by employing a cyclic feedback mechanism from previous time steps. Due to the sequential nature of user behavior data, LSTM can effectively capture their temporal evolution and dependencies. Each LSTM unit is composed of a memory cell, a hidden and three gating mechanism: input, output, and forget gate. The input gate i_t, forget gate f_t, output gate o_t, memory cell c_t and hidden state h_t at step t are computed as follows:

$$\begin{aligned}
i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \\
f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \\
o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \\
c_t &= f_t \odot c_{t-1} + i_t \odot tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \\
h_t &= o_t \odot tanh(c_t)
\end{aligned} \tag{1}$$

Here, σ is the logistic sigmoid function, $\tanh$ is the hyperbolic tangent function, $\odot$ denotes the element wise multiplication, and t is the time step for each individual session.

We implement a two-layer LSTM where the input is a sequence of user activity feature vectors of the sessions preceding the current session of interest. We term these preceding sessions utilized to model the ephemerality in user behavior as *ephemeral sessions*. In our implementation, we limit the number of ephemeral sessions to a maximum of five within the last four hours. We have varied both the number of ephemeral sessions and ephemeral time-window but observed no significant performance improvement in our joint modeling approach. In the case of less than five ephemeral sessions within four hours, we perform zero paddings. We show the ephemeral module in the bottom-right part of Figure 4, where the ephemeral LSTM iterates through the input sequence for five time-steps (each time step corresponds to one session). The final hidden state output from the ephemeral module can be considered as a latent representation of a user's recent activity summary.

Cyclic Module. Understanding the effect of time has been critical for effective user behavior modeling (Jing and Smola 2017; Beutel et al. 2018). However, utilizing temporal dynamics in a user-agnostic manner cannot exploit the individual user level cyclicity. To accommodate personalized temporal preference, the traditional approaches (e.g., probabilistic models) would require distinct parameters for each user, which is not practically feasible to learn or maintain (Kurashima, Althoff, and Leskovec 2018). In this regard, we argue that a user's time-specific historical activity can be leveraged to capture personalized temporal preference. To complement the recurrent network used for the ephemeral module and to facilitate end-to-end training, we propose to utilize a recurrent network for this purpose. The intuition behind this simplistic approach is to capture a user's past activity history at a specific time of the day into a latent representation by iterating through these activity sequences. As the recurrent networks employ cyclic feedback mechanism to update the current hidden states based on both current input and past hidden states, it is inherently suitable to aggregate overall cyclic history in the final embedding. For instance, for a user who typically engages in a long session at 5:00 pm during his commute, this approach makes it possible to use this knowledge for prediction.

Similar to the ephemeral module, we implement a two-layer LSTM for the cyclic module. However, contrary to the ephemeral module, here we use activities in a particular

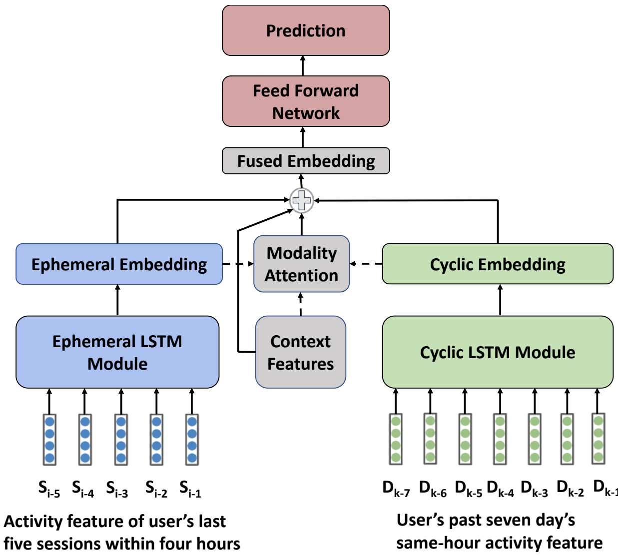


Figure 4: Proposed model architecture.

time-frame (i.e., "hour of the day", "day of the week") in the past few days or weeks to learn user's activity preference in that corresponding time-frame. Previous works have modeled temporal dynamics as a function of "hour of the day" to capture daily patterns (Kurashima, Althoff, and Leskovec 2018), and as a function of "hour of the day & day of the week (hour-weekday)" to capture weekly patterns (Jing and Smola 2017). In our empirical observation, we have observed both daily and weekly patterns. Subsequently, we have experimented with both scenarios to find the superiority of modeling daily patterns. We average user activities in each hour in the prior seven days to create an input sequence of length seven for the cyclic LSTM. We term each of these hourly averaged feature vectors as *cyclic sessions*. We show the cyclic module in the bottom-left part of Figure 4, where the cyclic LSTM iterates through the input sequence for seven time-steps (each time step corresponds to one hourly averaged activity features). The final hidden state output from the cyclic module can be considered as a latent representation of a user's cyclic activity summary in a particular time-frame.

Context Factors. Inspired by the previous successful use-cases of contextual factors (i.e., location, device, software client, or web browser for YouTube video recommendation (Beutel et al. 2018; Smirnova and Vasile 2017)), we utilize four contextual factors related to Snapchat sessions, which are: (1) *device connectivity*, whether the user is using WiFi connection or not; (2) *travel mode*, if the user is using the app in travel mode or not; (3) *app open state*, whether a user is opening the app by clicking on an app notification or not; (4) *app status*, whether the app was running in the background or not. We merge this context information in the one-hot encoded format with the output hidden embeddings from the ephemeral and cyclic module before feeding into a feed-forward neural network. Although these contextual factors are specific to Snapchat, similar contextual factors are prevalent across other social platforms. Moreover, other contextual factors inherent to a particular platform can be readily utilized in our architecture.

Modality Attention. The most intuitive and simplistic approach to aggregate multi-modal information is the naive concatenation of each modality (Moon, Neves, and Carvalho 2018). However, this naive concatenation of ephemeral and cyclic embedding treats both with equal importance in all instances. Nevertheless, in practicality, one modality can be more informative than the other. For example, while predicting a session, the user may not have any prior Snapchat sessions within the last four hours or any historical activities in the last seven days in that hour. Therefore, we employ a generalized modality attention module to attenuate or prioritize each modality for each prediction instance adaptively. This attention mechanism also enables us to quantitatively gauge each modality's impact on predictive modeling across numerous dimensions, i.e., user cohort, time, etc. (see experiments in section 6.5). We feed in the ephemeral (e), cyclic (c), and context (o) vectors into the attention module as input to generate a soft-attended attention vector for each modality $v \in \{e, c, o\}$ calculated as Equation 2:

$$\alpha_v = \frac{\exp\left(\phi\left(\mathbf{e}^v\right)\right)}{\sum_{\nu \in \{e,c,o\}} \exp\left(\phi\left(\mathbf{e}^\nu\right)\right)}, \tag{2}$$

where, $\mathbf{e}^v$ is the embedding vector of modality v, and $\phi(\cdot)$ is a mapping function implemented as a feed-forward neural network. Finally, we pass the fused embedding vectors through two feed forward neural networks before applying softmax to obtain a final prediction. In center of Figure 4, we show the modality attention fusing ephemeral, cyclic and contextual information. In coherence with the *attention* based joint modeling of *cyclicity* and *ephemerality*, we name our model as `CEAM`: Cyclic Ephemeral Attention Model.

6 Experiments

In this section, we evaluate the predictive performance of `CEAM` using two user activity datasets from Snapchat. We aim to answer the following experimental questions:

- **EQ1**: Can `CEAM` outperform existing methods in predicting user behavior?
- **EQ2**: How does the ephemeral and cyclic module in `CEAM` affect performance?
- **EQ3**: Can `CEAM` effectively model the cyclicity and ephemerality in user behavior?

Datasets and Experimental Setup

We perform experimental evaluations using two datasets collected from Snapchat. As mentioned previously (in Section 4.1), one dataset contains anonymized user activity data for a set of 20K randomly sampled monthly active users who were active at least once in each month from January 6, 2020 to February 23, 2020 (`MAU` dataset). We also extract a second dataset for 20K randomly sampled users who were active at least once each day within the aforementioned period (`DAU` dataset). In both cases, we collect 37 relevant user activity features for each session augmenting the ones previous studies on Snapchat (Tang et al. 2020; Yang et al. 2018; Liu et al. 2019). We perform min-max normalization (Patro and Sahu 2015) of each feature independently before training and testing. Both datasets span over seven weeks. We use the first four weeks for training purposes and the subsequent three weeks for testing.

Prediction Tasks. We define four specific user behavior prediction tasks for the upcoming session, which are the following:

- Task 1: Amount of time user will spend in the session.
- Task 2: Number of viewed discover stories.
- Task 3: User engagement with subscription content.
- Task 4: User engagement with recommended content.

Following Kooti et al.'s (Kooti et al. 2017) work on similar behavior prediction in Facebook, we frame these prediction tasks as classification problems by categorizing each behavior into multiple classes. For the first two tasks, we categorize the activity propensity into three classes: low, medium, and high, proportional to activity level (session length and discover view count). In both cases, the thresholds were selected to maintain roughly equal class balance. However, for the latter two tasks, we employ a binary classification scenario and predict whether the user engaged with the particular content category or not. We note that these user activities are also common on other social platforms such as Instagram and Tiktok, who display individual stories similar to Snapchat and recommended and subscription-based content. Hence, these prediction objectives can be easily transferred onto other social platforms.

Evaluation Metrics. We use *macro f1 score* as our main performance evaluation metrics. We run each training and testing experiments ten times, and report the average.

Compared Methods

We compare the performance of our method against the following state-of-the art methods to validate the accuracy of our user behavior prediction.

- **Copy Model (CM) (Anderson et al. 2014):** Predicting the current user behavior the same as users' last session's. This can be deemed as the most naive version of ephemerality based prediction.
- **LSTM (Jing and Smola 2017):** LSTM has shown promising results in several sequential user behavior modeling tasks. We adopt the methods proposed in (Jing and Smola 2017). Similar to them, we generate embedding vectors for timing and interval of ephemeral sessions after passing through embedding layers and concatenate with the session activity features. We implement a two-layer LSTM that iterates over the ephemeral session's feature vectors and feeds the output hidden embedding into a two-layer fully connected network to generate the prediction.
- **TLSTM (Zhu et al. 2017):** In (Zhu et al. 2017), TLSTM has been introduced for user behavior modeling where the time interval between two consecutive actions has been used to moderate a gating mechanism to update the hidden states of LSTM for improved performance. We implement the architecture proposed in (Zhu et al. 2017), and feed

Table 1: Prediction performance (macro f1-score) of `CEAM` on all tasks and both datasets `MAU` and `DAU`.

		Task 1	Task 2	Task 3	Task 4
MAU	**CP**	.375±.000	.296±.000	.320±.000	.321±.000
	LSTM	.480±.002	.482±.001	.685±.000	.700±.000
	TLSTM	.470±.005	.472±.003	.680±.003	.695±.002
	CEAM	**.500±.001**	**.521±.001**	**.723±.000**	**.743±.000**
DAU	**CP**	.370±.000	.288±.000	.471±.000	.319±.000
	LSTM	.480±.001	.475±.002	.682±.000	.698±.000
	TLSTM	.478±.006	.469±.004	.680±.002	.694±.003
	CEAM	**.500±.002**	**.518±.002**	**.722±.000**	**.744±.000**

Table 2: Ablation studies to show the contribution to performance improvement (macro f1-score) by different modules.

		Task 1	Task 2	Task 3	Task 4
MAU	**Ephemeral**	.479±.002	.479±.002	.683±.000	.701±.000
	Cyclic	.459±.003	.480±.001	.673±.000	.700±.000
	Combined	.500±.002	.519±.001	.721±.000	.741±.000
	CEAM	**.501±.001**	**.521±.001**	**.723±.000**	**.743±.000**
DAU	**Ephemeral**	.478±.002	.474±.002	.680±.001	.694±.000
	Cyclic	.460±.002	.480±.001	.675±.000	.700±.000
	Combined	.499±.001	.517±.002	.720±.001	.742±.000
	CEAM	**.500±.002**	**.518±.002**	**.722±.000**	**.744±.000**

in the ephemeral sessions along with session interval of consecutive sessions.

`CEAM` has access to contextual information to enhance predictions. Although not utilized in the proposed LSTM (Jing and Smola 2017) and TLSTM (Zhu et al. 2017), for a fair comparison, in our implementation, we also include the contextual information in the fully connected layers similar to `CEAM`. Previously, latent variable models (e.g., Markov models, hidden Markov models) and Poisson process models have been used for similar user behavior modeling. However, several recent works employing LSTM based models have consistently and significantly outperformed the aforementioned approaches (Jing and Smola 2017; Kurashima, Althoff, and Leskovec 2018). Hence, we omit the comparison against these methods.

Model Implementations

We implement a two-layer LSTM network for the ephemeral and cyclic module, with an embedding size of 32 in both layers. We set the second layer's dropout rate to 0.5 and use ReLU as the activation function. We implement `CEAM` and other neural models using PyTorch[1]. All the models are optimized using Adam Algorithm (Boyd, Parikh, and Chu 2011), with an initial learning rate of 0.001, and an L2 regularization of $1e-6$. We set the batch size to 512. We train all models to a maximum of 50 epochs with early stopping on the validation set. All the hyper-parameters were selected empirically using a grid search on a held-out validation set.

Prediction Performance

To answer the first experimental question, we report the prediction performance *(macro f1-score)* of `CEAM` along with the compared methods for all four tasks on both datasets in Table 1. We observe that `CEAM` outperforms all the other methods in all four tasks in both datasets. We note that both `LSTM` and `tLSTM` perform on a similar level. However, `CEAM` outperforms both these models by at most 10%. Although both `LSTM` and `tLSTM` uses session timing as contextual information, they significantly underperform `CEAM`, which validates our argument to model cyclicity at the individual user level.

Ablation Study

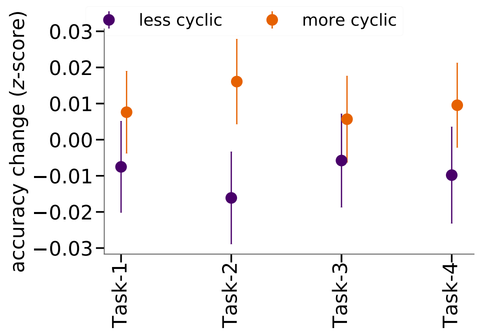
To answer the second experimental question, we perform several ablation studies by developing three variations of the proposed model. (1) `EPHEMERAL`: In the first variation, we use only the ephemeral module and feed the output embedding concatenated with the contextual embedding into a two-layer fully connected network for prediction generation. In contrast with baseline `LSTM`, we do not utilize session-timing or session interval as features in the ephemeral module. (2) `CYCLIC`: Next, we use only the cyclic module in a similar fashion (3) `COMBINED`: Then, we employ both cyclic and ephemeral module and concatenate their output embedding along with contextual embedding before feeding into the fully connected layers. In Table 2, we report the prediction performance for all variations. Here, we observe that both cyclic and ephemeral module are in general predictive across all tasks in both datasets. Their predictive accuracies are in a similar range, with `Ephemeral` being higher in majority of the cases. However, the combination of both shows consistent performance improvements than each employed individually. This further increases with the use of self-attention. We see at most 10% performance improvement by `CEAM` over single module approach. This improvement validates the presence of complementary information in both behavioral dynamics and the necessity of joint modeling.

Model Sanity Check

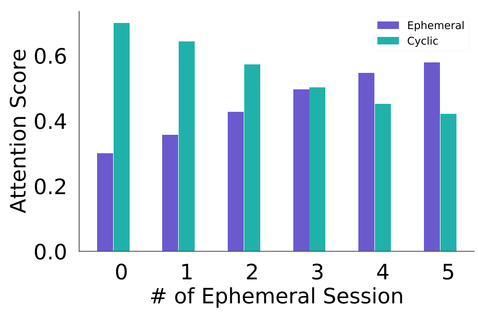
To answer the third experimental question, we design multiple experiments with the goal to understand *for whom, how, and when* `CEAM` improves prediction performance. We perform these experiments on the `MAU` dataset. First, we investigate *for whom*, `CEAM` shows better performance. The underlying motivation behind modeling cyclicity on a per-user basis is to capture an individual user's app usage regularities for improved predictability. Therefore, we would expect `CEAM` to be more effective for more cyclic users. To quantify the performance improvement in relation to a user's cyclicity, we separate users into two groups based on a median split on *cyclicity scores* (defined in Section 4). One group consists of the top 50% cyclic users *(more-cyclic)*, and the rest were in the other *(less-cyclic)*.

We introduce the `CYCLIC` module in predictions, and then calculate the change in accuracy for the user's prediction in both the groups.We transform the raw accuracy change values into z-score to reduce sensitivity to inconsistent magnitudes of absolute values (Golder and Macy 2011), used in prior related works (Saha and De Choudhury 2017). By definition, z-score represents the distance between raw value and population mean in units of standard deviation (Golder and Macy 2011). In Figure 5a, we report the mean z-score

[1] www.pytorch.org/



(a)

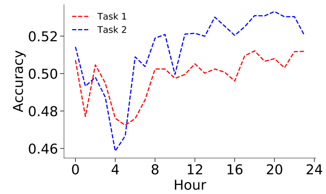


(b)

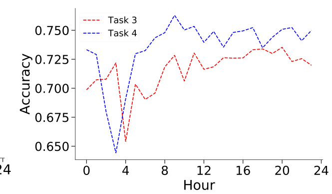
Figure 5: (a) Accuracy change (z-transformed, 0 shows mean improvement) for more-cyclic and less-cyclic user groups for all four tasks after adding the cyclic module. We gain greater improvement for more-cyclic users. (b) The variation in ephemeral and cyclic attention scores as the ephemeral session number varies. The cyclic attention is greater in fewer ephemeral sessions, and the ephemeral attention increases along with the increase in ephemeral session number.

of accuracy change for both user groups for all four tasks. Here, zero z-score refers to the mean accuracy improvement for all the users (population) over the ephemeral module only performance, which is $\approx$10%. Above zero z-score indicates greater improvement than the overall population mean improvement and vice versa. We observe that in all the four tasks, *more-cyclic* users show greater accuracy improvement than the population mean whereas *less-cyclic* users show less accuracy improvement, which resonates with our initial intuition about the model.

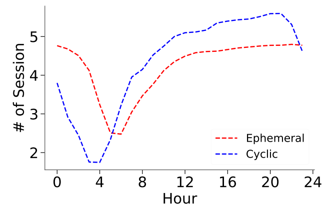
Next, we investigate *how* CEAM better exploits both cyclic and ephemeral aspects of human behavior. In CEAM, we employ a self-attention mechanism to fuse information from both these modalities. We ask whether the attention module can adaptively prioritize one over the other depending on the information of each modality. To quantify dynamic preference, we calculate the shift in attention in relation to the number of available sessions in each module. As example, in Figure 5b, we show the variation of ephemeral and cyclic attention weight as the number of ephemeral sessions vary. Following the intuition, the cyclic module gets greater attention in case of fewer ephemeral sessions. And as the ephemeral session number increases, so does the attention



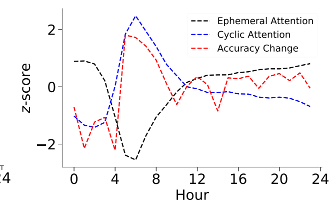
(a)



(b)



(c)



(d)

Figure 6: The hourly averaged accuracy for all four tasks (a) Task 1 and Task 2, (b) Task 3 and Task 4. (c) The average number of ephemeral and cyclic sessions in each hour. (d) Hourly average ephemeral and cyclic attention weight, and accuracy change after incorporating cyclic module (z-transformed).

weight for the ephemeral module.

Lastly, we explore the temporal aspect of prediction performance. In Figure 6a and 6b, we show the prediction accuracy over hour of the day for Task-1&2 and Task-3&4 respectively. We observe that, for all four tasks, the accuracy remains similar from around 8 AM till midnight. However, there are sharps drops between midnight to around 6 AM in the morning. To better understand it, we plot the average number of ephemeral and cyclic sessions across the day in Figure 6c. We notice that both ephemeral and cyclic session average drops during late night to early morning compared to the rest of the day, which correlates with the lower accuracy period. Therefore, we can reasonably attribute the lower accuracy period to the shortage of cyclic and ephemeral activities around that time.

We further explore how CEAM exploits information from both modalities to improve prediction across hours. In Figure 6d, we plot the hourly average attention weight for the ephemeral and cyclic module along with hourly average accuracy change after adding the cyclic module. Again, we utilize the z-transformation for normalized comparison across several measures. We observe a drop in ephemeral weight and increment in both cyclic weight and accuracy change between 4 AM to 8 AM. The positive increase in cyclicity indicates that CEAM relying more on cyclic information around that time period when there are fewer ephemeral sessions, as shown in Figure 6c. And, due to this adaptive and better utilization of cyclicity, we improve prediction.

7 Discussion

Model Robustness

User agnostic. CEAM does not require any user specific information and does not learn any user-specific parameters. Hence, our proposed model can be readily used for any new

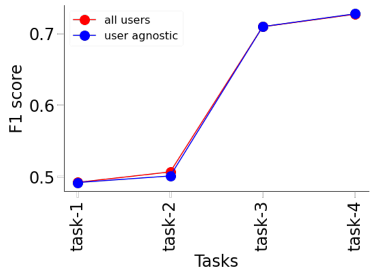


Figure 7: F1-score for user-agnostic vs. non user-agnostic training and testing.

users in cold-start situations (Schein et al. 2002). To drive this point home, we simulated a similar scenario, where we only evaluate for users that do not have sessions in the training set. In Figure 7, we show the prediction F1-score for the user-agnostic and non user-agnostic case. We observed similar performance in both cases, which shows the usability of our method in cold-start situations.

Ethical and Privacy Implications

Several recent data breaches and reported misuses of sensitive and private user data, have resulted in growing public concerns and increases in regulations (Cadwalladr and Graham-Harrison 2018). Any experimental study dealing with potentially sensitive data demands a statement regarding ethical conduct and secure data handling. In our study, user activity data from the social app Snapchat was used for empirical and experimental purposes. Following Snapchat's in-house commitment towards user-privacy and data-protection policies, our dataset was anonymized before our analyses and was void of any form of personally identifiable demographic or typographic information. Moreover, the user activity data was only comprised of amounts and frequencies of activities: no information such as communication content, type of content, or source of viewed content were used in any part of this study. All the experiments were conducted within Snapchat's internal secure storage systems, and data was not stored on local computers or outside the Snap Inc. ecosystem.

Given that our models well predicted user behavior without using personally identifiable information and user attributes, this modeling approach is applicable to several scenarios. First, because our behavior modeling does not use any identifiable user attributes, it may be preferable in highly sensitive settings. Second, because the model uses no platform-specific user attributes, it would be relatively straightforward to apply this technique to other platforms. This could be useful to technology designers (e.g. for modeling the behavior of users in other mobile apps) or to online communities that stand to benefit from user behavior modeling.

Limitations and Future Works

Our proposed method utilizes only the propensity of user actions and activities while ignoring their qualitative aspects. For example, we only consider the time spent viewing multimedia content, but we disregard the nature of the content being watched. In future works, the observed temporal activity variations can be further explored to learn each user's preference towards certain content and ads across the day, which can help in identifying suitable time-periods to distribute content and ads. Better personalized content allocation without compromising privacy can benefit the 100B social platform industry and the platforms' users alike. Moreover, our flexible modeling approach is suitable for integrating a multitude of contextual factors (e.g., seasonality, weather, location) when available, depending on the social platform, to improve the prediction performance. Additionally, further experiments can be conducted to examine how the time difference between consecutive sessions affects the ephemerality aspect of user behavior. Future work might explore an online training setting in which user behavioral modeling is integrated with a real-time prediction mechanism.

8 Conclusion

In this paper, we explored one way to improve user behavior modeling on online social platforms by including cyclical behavior in the prediction. Using Snapchat data, we demonstrated regularities in people's behavior, both at the collective and the individual level. We then proposed an end-to-end neural framework that leverages both cyclic and ephemeral aspects of people's daily lives for improved prediction. Importantly, our method is *personalized, but agnostic* to privacy-invasive data: We do not use any user-typographic or demographic information, and we avoid any sort of long-term data based user profiling. Short-term data and dynamically generated behavioral features are a more ethical approach for responsible user modeling. We evaluated the efficacy of our method using four prediction tasks on two datasets from Snapchat. Our method outperforms existing methods by on average 7% (up to 10%) (macro f1-score). Empirically, we show that our model can successfully capture the cyclicity in individual user behavior. While our work focuses on Snapchat, the demonstrated insights and proposed modeling approach can potentially motivate similar explorations and modeling in other social platforms to increase business value and deliver a better user experience.

References

Adar, E.; Teevan, J.; and Dumais, S. T. 2008. Large scale analysis of web revisitation patterns. In *SIGCHI*, 1197–1206.

Anderson, A.; Kumar, R.; Tomkins, A.; and Vassilvitskii, S. 2014. The dynamics of repeat consumption. In *WWW*, 419–430.

Banovic, N.; Buzali, T.; Chevalier, F.; Mankoff, J.; and Dey, A. K. 2016. Modeling and understanding human routine behavior. In *SIGCHI*, 248–260.

Benson, A. R.; Kumar, R.; and Tomkins, A. 2016. Modeling user consumption sequences. In *WWW*.

Berkovsky, S.; Kuflik, T.; and Ricci, F. 2008. Mediation of user models for enhanced personalization in recommender systems. *User Modeling and User-Adapted Interaction* .

Beutel, A.; Covington, P.; Jain, S.; Xu, C.; Li, J.; Gatto, V.;

and Chi, E. H. 2018. Latent cross: Making use of context in recurrent recommender systems. In *WSDM*, 46–54.

Boyd, S.; Parikh, N.; and Chu, E. 2011. *Distributed optimization and statistical learning via the alternating direction method of multipliers*. Now Publishers Inc.

Cadwalladr, C.; and Graham-Harrison, E. 2018. Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian* 17.

Chen, C.; Kim, S.; Bui, H.; Rossi, R.; Koh, E.; Kveton, B.; and Bunescu, R. 2018. Predictive analysis by leveraging temporal user behavior and user embeddings. In *CIKM*, 2175–2182.

Cheng, J.; Bernstein, M.; Danescu-Niculescu-Mizil, C.; and Leskovec, J. 2017. Anyone can become a troll: Causes of trolling behavior in online discussions. In *CSCW*, 1217–1230.

Culotta, A. 2014. Estimating county health statistics with Twitter. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1335–1344. ACM.

Das Swain, V.; Saha, K.; Rajvanshy, H.; Sirigiri, A.; Gregg, J. M.; Lin, S.; Martinez, G. J.; Mattingly, S. M.; Mirjafari, S.; Mulukutla, R.; et al. 2019. A Multisensor Person-Centered Approach to Understand the Role of Daily Activities in Job Performance with Organizational Personas. *Proc. IMWUT* .

Golder, S. A.; and Macy, M. W. 2011. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science* 333(6051): 1878–1881.

Golder, S. A.; Wilkinson, D. M.; and Huberman, B. A. 2007. Rhythms of social interaction: Messaging within a massive online network. In *Communities and technologies 2007*, 41–66. Springer.

Grinberg, N.; Naaman, M.; Shaw, B.; and Lotan, G. 2013. Extracting Diurnal Patterns of Real World Activity from Social Media.

Habib, H.; Shah, N.; and Vaish, R. 2019. Impact of Contextual Factors on Snapchat Public Sharing. In *SIGCHI*, 1–13.

Hidasi, B.; Karatzoglou, A.; Baltrunas, L.; and Tikk, D. 2015. Session-based recommendations with recurrent neural networks. *arXiv preprint arXiv:1511.06939* .

Hidasi, B.; Quadrana, M.; Karatzoglou, A.; and Tikk, D. 2016. Parallel recurrent neural network architectures for feature-rich session-based recommendations. In *RecSys*, 241–248.

Jing, H.; and Smola, A. J. 2017. Neural survival recommender. In *WSDM*, 515–524.

Kapoor, K.; Kumar, V.; Terveen, L.; Konstan, J. A.; and Schrater, P. 2015a. " I like to explore sometimes" Adapting to Dynamic User Novelty Preferences. In *RecSys*, 19–26.

Kapoor, K.; Subbian, K.; Srivastava, J.; and Schrater, P. 2015b. Just in time recommendations: Modeling the dynamics of boredom in activity streams. In *WSDM*, 233–242.

Kawale, J.; Pal, A.; and Srivastava, J. 2009. Churn prediction in MMORPGs: A social influence based approach. In *2009 International Conference on Computational Science and Engineering*, volume 4, 423–428. IEEE.

Kooti, F.; Aiello, L. M.; Grbovic, M.; Lerman, K.; and Mantrach, A. 2015. Evolution of conversations in the age of email overload. In *WWW*, 603–613.

Kooti, F.; Subbian, K.; Mason, W.; Adamic, L.; and Lerman, K. 2017. Understanding short-term changes in online activity sessions. In *WWW Companion*, 555–563.

Koren, Y. 2009. Collaborative filtering with temporal dynamics. In *KDD*.

Kurashima, T.; Althoff, T.; and Leskovec, J. 2018. Modeling interdependent and periodic real-world action sequences. In *WWW*.

Liu, Y.; Shi, X.; Pierce, L.; and Ren, X. 2019. Characterizing and Forecasting User Engagement with In-app Action Graph: A Case Study of Snapchat. In *KDD*, 2023–2031.

Lo, C.; Frankowski, D.; and Leskovec, J. 2016. Understanding behaviors that lead to purchasing: A case study of pinterest. In *KDD*.

Moon, S.; Neves, L.; and Carvalho, V. 2018. Multimodal named entity recognition for short social media posts. *arXiv preprint arXiv:1802.07862* .

Murnane, E. L.; Abdullah, S.; Matthews, M.; Choudhury, T.; and Gay, G. 2015. Social (media) jet lag: How usage of social technology can modulate and reflect circadian rhythms. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 843–854.

Patro, S.; and Sahu, K. K. 2015. Normalization: A preprocessing stage. *arXiv preprint arXiv:1503.06462* .

Pierson, E.; Althoff, T.; and Leskovec, J. 2018. Modeling individual cyclic variation in human behavior. In *WWW*, 107–116.

Saha, K.; and De Choudhury, M. 2017. Modeling stress with social media around incidents of gun violence on college campuses. *Proceedings of the ACM on Human-Computer Interaction* 1(CSCW): 92.

Saha, K.; Grover, T.; Mattingly, S.; Das Swain, V.; Gupta, P.; Martinez, G. J.; Robles-Granda, P.; Mark, G.; Striegel, A.; and De Choudhury, M. 2021a. Person-Centered Predictions of Psychological Constructs with Social Media Contextualized by Multimodal Sensing. *PACM IMWUT* .

Saha, K.; Liu, Y.; Vincent, N.; Chowdhury, F. A.; Neves, L.; Shah, N.; and Bos, M. W. 2021b. AdverTiming Matters: Examining User Ad Consumption for Effective Ad Allocations on Social Media. In *CHI*.

Satopaa, V.; Albrecht, J.; Irwin, D.; and Raghavan, B. 2011. Finding a" kneedle" in a haystack: Detecting knee points in system behavior. In *2011 31st international conference on distributed computing systems workshops*, 166–171. IEEE.

Schein, A. I.; Popescul, A.; Ungar, L. H.; and Pennock, D. M. 2002. Methods and metrics for cold-start recommendations. In *SIGIR*, 253–260.

Smirnova, E.; and Vasile, F. 2017. Contextual sequence modeling for recommendation with recurrent neural networks. In *Proceedings of the 2nd Workshop on Deep Learning for Recommender Systems*, 2–9.

Statista. 2020. https://www.statista.com/statistics/545967/snapchat-app-dau/.

Tang, X.; Liu, Y.; Shah, N.; Shi, X.; Mitra, P.; and Wang, S. 2020. Knowing your FATE: Friendship, Action and Temporal Explanations for User Engagement Prediction on Social Apps. *arXiv preprint arXiv:2006.06427* .

Teevan, J.; Adar, E.; Jones, R.; and Potts, M. 2006. History repeats itself: repeat queries in Yahoo's logs. In *SIGIR*, 703–704.

Teevan, J.; Adar, E.; Jones, R.; and Potts, M. A. 2007. Information re-retrieval: repeat queries in Yahoo's logs. In *SIGIR*, 151–158.

Trouleau, W.; Ashkan, A.; Ding, W.; and Eriksson, B. 2016. Just one more: Modeling binge watching behavior. In *KDD*, 1215–1224.

Verbeke, W.; Martens, D.; and Baesens, B. 2014. Social network analysis for customer churn prediction. *Applied Soft Computing* 14: 431–446.

Yang, C.; Shi, X.; Jie, L.; and Han, J. 2018. I Know You'll Be Back: Interpretable New User Clustering and Churn Prediction on a Mobile Social Application. In *KDD*, 914–922.

Zhang, B.; Kreitz, G.; Isaksson, M.; Ubillos, J.; Urdaneta, G.; Pouwelse, J. A.; and Epema, D. 2013. Understanding user behavior in spotify. In *IEEE INFOCOM*, 220–224. IEEE.

Zhu, Y.; Li, H.; Liao, Y.; Wang, B.; Guan, Z.; Liu, H.; and Cai, D. 2017. What to Do Next: Modeling User Behaviors by Time-LSTM. In *IJCAI*, volume 17, 3602–3608.